## 6.1 Assessment of satellite radiometry in the visible domain

F. Mélin, E.C. Joint Research Centre, Institute for Environment and Sustainability, Ispra 21027, Italy
B.A. Franz, Ocean Biology Processing Group, NASA, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

Abstract

Marine reflectance and chlorophyll-a concentration are listed among the Essential Climate Variables by the Global Climate Observing System. To contribute to climate research, the satellite ocean color data records resulting from successive missions need to be consistent and well characterized in terms of uncertainties. This chapter reviews various approaches that can be used for the assessment of satellite ocean color data. Good practices for validating satellite products with *in situ* data and the current status of validation results are illustrated. Model-based approaches and inter-comparison techniques can also contribute to characterize some components of the uncertainty budget, while time series analysis can detect issues with the instrument radiometric characterization and calibration. Satellite data from different missions should also provide a consistent picture in scales of variability, including seasonal and inter-annual signals. Eventually, the various assessment approaches should be combined to create a fully characterized climate data record from satellite ocean color.

Keywords: ocean color, marine reflectance, chlorophyll-a concentration, optical properties assessment, validation, uncertainties

### 1. Introduction

Standardization of methods to assess and assign quality metrics to satellite ocean color radiometry and derived geophysical products has become paramount with the inclusion of the marine reflectance and chlorophyll-a concentration as Essential Climate Variables (ECV; [1]) and the recognition that optical remote sensing of the oceans can only contribute to climate research if and when a continuous succession of satellite missions can be shown to collectively provide a consistent, long-term record with known uncertainties. In 20 years, the community has made significant advancements toward that objective, but providing a complete uncertainty budget for all products and for all conditions remains a daunting task. In the retrieval of water-leaving marine reflectance from observed top-of-atmosphere radiances, the sources of uncertainties include those associated with propagation of sensor noise and radiometric calibration and characterization errors, as well as a multitude of uncertainties associated with the modeling and removal of effects from the atmosphere and sea surface. This chapter describes some common approaches used to assess quality and consistency of ocean color satellite products and reviews the current status of uncertainty quantification in the field. Its focus is on the primary ocean color product, the spectrum of marine reflectance $R_{rs}$, but uncertainties in some derived products such as the chlorophyll-a concentration (Chla) or inherent optical properties (IOPs) will also be considered.

## 2. Validation of satellite products

The primary method to assess satellite data is through direct comparison of a satellite product with near contemporaneous and co-located *in situ* measurements of the same quantity. Using the field data as a reference, such comparisons can provide estimates of the uncertainty associated with the satellite product. For derived products that rely on empirical algorithms, the *in situ* validation dataset should be independent of any measurements used to define or tune the satellite retrieval algorithm. Unfortunately, the collection of high quality field measurements of optical radiometry for validation is challenging due to difficult environmental conditions, cloud cover and other factors that restrict remote observation, and logistical difficulties of ocean access, thus leading to a relatively limited geographic and temporal sampling of available *in situ* validation data. As an alternative, simulated data sets have been used for validation studies, since they can be considered error-free and can cover a large range of optical conditions [2]. They can also be produced at any desired wavelengths while validation of multi-spectral quantities with field data may be hampered by differences in wavelengths between the quantities to be compared. These differences need to be considered and possibly corrected in the validation exercise ~~and possibly corrected for~~ (see Section 3.1). This Section focuses on assessment of satellite product uncertainties through comparisons with field observations, including a description of validation protocols and metrics and a discussion on validation results. Error propagation techniques and use of atmospheric correction and bio-optical models to assess confidence intervals are also briefly reviewed.

### 2.1 Validation protocol

The validation protocol needs to be well documented to ensure consistency of approach between missions and products and reproducibility between studies. The first step is the construction of the validation, or match-up, data set. A match-up refers to the meaningful association of a satellite value with its counterpart from field observations. This entails the extraction from the overall satellite record of a subset of pixels or grid points, usually a square of $N_s x N_s$ elements centered on the location of the field value and separated in time by less than a small interval $\Delta t$. From the extracted values, one can derive three main statistics: i) the fraction *fv* of valid retrievals among the $N_s x N_s$ potential values, ii) the average (or median) satellite value, and iii) the spatial coefficient of variation ($CV_s$), which is the ratio of the standard deviation within the $N_s x N_s$ valid satellite measurements and the average value. A high $CV_s$ means that the satellite retrievals show a large heterogeneity, which in turn suggests a reduced probability that the *in situ* point measurement is representative of the region observed by the satellite. In the interval $\pm\Delta t$, there might be $N_t$ field observations collected, so that an average (or median) value and a temporal $CV_t$ can also be calculated. In that case, a high CV is indicative of changing conditions at the location of the measurements. Eventually, the match-up selection protocol defines the allowed values for maximum $\Delta t$, minimum *fv* and maximum $CV_s$, as well as minimum $N_t$ and

maximum $CV_t$ if applicable. Then, the satellite average (or median) value can be compared with the average field observation or its datum closest to the satellite overpass time.

The choice of the threshold values should allow for a sufficient number of match-ups to conduct a proper statistical analysis while maintaining the validity of the comparison. This compromise should take into account the expected environmental conditions. For instance Bailey and Werdell [3] have selected $N_s$=5, $\Delta$t=3-h, *fv* of 50% and $CV_s$ of 15% for a global validation analysis that relies on many points in open ocean where conditions are thought more stable. Zibordi et al. [4] have used $N_s$=3, $\Delta$t=2-h, *fv* of 100% and $CV_s$ of 20% for validation at a coastal station, the Acqua Alta Oceanographic Tower (AAOT) located in the northern Adriatic Sea. Recommended values can be given, with $N_s$ of 3 to 5, $\Delta$t of 1 to 4-h, *fv* larger than 50%, and $CV_s$ of less than 20% for some of the products being validated (typically ~~for~~ $R_{rs}$ at a selected wavelength). The choice of threshold values should be adapted to the conditions associated with the validation exercise, with dynamic coastal environments generally requiring more stringent criteria. It is good practice to test several thresholds to assess how validation statistics are affected. For instance, Feng et al. [5] showed how validation statistics improved with more stringent match-up selection criteria. Such an analysis can also provide insight into the degree of representativeness of the comparison, quantifying the discrepancy in scale and time of observation between the two measurement systems. For $R_{rs}$ validation, it is also recommended to operate the selection protocol on the spectrum as a whole and not independently on separate bands; indeed the selection of varying numbers of data points for the different channels hinders a consistent assessment over the spectral domain of interest.

Ideally, a validation analysis should ~~be based on~~integrate the knowledge of the uncertainties associated with the field observations. Comprehensive validation exercises often combine *in situ* data collected by a variety of disparate systems and investigators using different instruments and measurement techniques. In such cases it is recommended to assess the dependence of the validation results on the different data sets gathered for the exercise. More generally, if enough *in situ* data are available, the sensitivity of the validation statistics to particular sets of match-ups can be quantified by bootstrapping techniques [6].

Figure 6.1 shows an example of match-ups obtained at the AAOT site for the sensor MODIS-Aqua, both in terms of reflectance $R_{rs}$ and aerosol optical thickness $\tau_A$. Over the period 2002-2012, the number of match-ups found for $R_{rs}$ is 549, with $N_s$=3, $\Delta$t=1h, and $CV_s$=20% for $R_{rs}$ between 488 and 547 nm. More match-ups are obtained for the aerosol products, with the added conditions of $N_t$=2 and $CV_t$=20% for $\tau_A$ at 488 nm.

## 2.2 Validation metrics

There are a host of statistical metrics that can be used to compare two data sets, but a minimum set for validation would include the number of match-ups (together with the number of potential

match-ups) and estimates of the scatter and systematic difference (bias) between the two distributions. According to the range of values considered, these statistics can be expressed as absolute or relative values, with a prior log transformation typically applied for Chla or IOPs. For radiometric products, it is important to document both a measure of uncertainty in radiometric units ($sr^{-1}$ for $R_{rs}$) and a measure of relative uncertainty. Indeed, relative differences tend to increase when the values of $R_{rs}$ are small, up to tens of percent if the *in situ* value is near zero. In that case, the difference in radiometric units is more meaningful.

Relative differences between satellite products $(y_i)_{i=1,N}$ and field observations $(x_i)_{i=1,N}$ can be expressed in %, and computed in terms of mean absolute difference or mean difference (i.e., bias) with respect to the field observations:

$$(1) \quad |\psi| = 100.\frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - x_i|}{x_i}$$

$$(2) \quad \psi = 100.\frac{1}{N}\sum_{i=1}^{N}\frac{y_i - x_i}{x_i}$$

while the equivalent metrics in geophysical units can be computed as:

$$(3) \quad |\delta| = \frac{1}{N}\sum_{i=1}^{N}|y_i - x_i|$$

$$(4) \quad \delta = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i) = \bar{y} - \bar{x}$$

where the overbar means average values. Root-mean-square differences between the satellite and *in situ* measurements can be written as:

$$(5) \quad \Delta = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2}$$

$$(6) \quad \Delta_u = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y} - x_i + \bar{x})^2} = \sqrt{\Delta^2 - \delta^2}$$

The total root-mean-square difference Δ can be partitioned into a part due to the bias δ and the unbiased (or centered) root-mean-square difference $\Delta_u$ quantifying non-systematic effects. In the above equation, the summation operator has been used (which means that quantities are averages), but other operators can be ~~used~~ preferred like the median~~,~~ or some form of inter-quantile statistics. Other metrics can be included like the coefficient of determination, $r^2$, slope and intercept of linear regression, average ratios, etc.~~s…~~ For spectral quantities like $R_{rs}$ input to bio-optical algorithms, quantifying how well the spectral shape is respected by the satellite

products is also worthy information, that can be quantified by the $\chi^2$ distribution measuring the goodness of fit between *in situ* and satellite $R_{rs}$ normalized at one wavelength of reference [7].

To document differences (including their systematic component), it is recommended to compute at least $|\psi|$, $\psi$, $\Delta$ (or $\Delta_u$) and $\delta$. As illustration, Figure 6.2 shows spectral target diagrams that display $\Delta_u$ and $\delta$ simultaneously for various sensors and two validation sites, AAOT and the Marine Optical Buoy (MOBy) near Hawaii [8]. By construction, $\Delta$ is the distance between a point and the origin (6). The value of $\Delta$ (or $\Delta_u$) decreases with increasing wavelength for MODIS compared to MOBy data, with virtually no bias, which is expected considering the role of that site for vicarious calibration [9]. This is also true for the validation results obtained at AAOT, but significant values can be observed for the bias $\delta$, generally negative for MODIS, and positive for MERIS and SeaWiFS (validation results for MODIS are the same as shown on Figure 6.1).

2.3 Analysis of validation results

Performing accurate radiometric *in situ* oceanographic measurements to derive $R_{rs}$ is difficult and expensive, with the implication that match-ups with satellite data are not abundant and are unevenly distributed in space and time [3]. More match-ups are available for Chla, although large expanses of ocean remain devoid of validation data [10]. Using the SeaWiFS Bio-optical Archive and Storage system (SeaBASS) [11], a global community field data repository for marine bio-optical measurements, and following the standard protocols in [3], one obtains less than 1000 match-ups for the 13-year SeaWiFS mission (numbers vary by wavelength).

Figure 6.3 is a snapshot of validation results expressed as spectra of $\Delta$ between satellite and field data. MODIS-Aqua $\Delta$ values are illustrated for various data sets (Figure 6.3a), SeaBASS, BiOMaP (representative of European waters [12]), and MOBy, as well as various AERONET-OC sites [13] located in the northern Adriatic Sea (AAOT), Baltic Sea (Gustaf Dalen, GDLT, and Helsinki, HLT, ~~lighthouses~~light towers), Black Sea (Gloria, GLR), Chesapeake Bay entrance (COVE), coastal Gulf of Mexico (WAVE), coastal southern California (USC), and Persian Gulf (Abu al-Bukhoosh Platform, AABP). The number of match-ups varies from 15 (AABP) to 549 (AAOT); the number of wavelengths represented is also variable (e.g., the SeaBASS validation results are shown at 412, 443, 488 and 667 nm only). Some spectra show a value at 547 and 555 nm, the latter being a band not originally intended for ocean color applications by the MODIS mission. Most $\Delta$ values are found between 0.0008 and 0.0015 $sr^{-1}$ at 412 nm, down to between 0.0002 to 0.0004 $sr^{-1}$ at 667 nm (with the exception of MOBy where $\Delta$ is lower). The $\Delta$ values are partly conditioned by the actual $R_{rs}$ values; for instance, $\Delta$ is lowest at MOBy in the green bands, whereas it is lowest at the Baltic sites in the blue part of the spectrum, where $R_{rs}$ is often very low.

There is a clearer consistency of Δ spectra when considering validation results at a single site for different missions processed with the same NASA-standard algorithms (Figure 6.3b at AAOT). Results for standard ESA MERIS products are shown on Figure 6.3c, for the Bohai Sea (N=17) [14], the northwest Mediterranean Sea (N=64 except at 412 nm) [15], South African coastal waters (N=14) [16], and AERONET-OC sites in the northern Adriatic Sea (AAOT, N=86), Baltic Sea (GDLT and HLT, N=39) and Black Sea (GLR, N=12) [17]. For completeness, Global Imager (GLI) results are also shown (N=435 at 443 nm) [18]. MERIS Δ values tend to be fairly high, particularly in the blue. The case of the Bohai Sea is fairly unique and associated with highly scattering waters with $R_{rs}$ maxima beyond 550 nm [14]. This type of Δ spectra should be confirmed with more match-ups.

Finally, interesting studies have been done to compare atmospheric correction schemes with the same validation data set [7]. Two examples of such exercises are reported on Figure 6.3d, comparing SeaDAS results with other schemes [19,20]. Validation statistics appear fairly consistent for a given sensor, with the Δ values associated with the standard scheme often being the lowest. The family of Δ curves of Figure 6.3 could be presented for other statistical indicators. Relative differences, $|\psi|$ or $\psi$, would show more variations particularly between different locations, $|\psi|$ varying from 10% to tens of percent. In fact, $|\psi|$ spectra are often an inverted image of $R_{rs}$ spectra, with $|\psi|$ values that are high in red bands in oligotrophic waters or that may exceed 100% at 412 nm in absorbing waters like in the Baltic Sea [4].

Other $R_{rs}$ validation exercises have of course been conducted, applied to specific sensors like OCTS [21],VIIRS [22], or GOCI [23,24], specific coastal regions (e.g., coastal Chinese waters [25]), or to test alternative atmospheric corrections [26,27]. Many of these studies suffer from a limited number of match-ups often collected in restricted geographical areas. More work is needed to extend, analyze and understand validation results across missions, atmospheric correction schemes, and field data sets or locations, with the goal to enable the extension of point validation results sparsely distributed in space to the global ocean. This issue will be further discussed in this and the following sections.

Validation analyses ideally should go beyond simply providing statistics for a given location and/or season by investigating possible dependences of the validation results on time or season, geometry of observation and illumination, atmospheric conditions, or marine properties. The benefit can be two-fold, as such studies can provide insights into the reasons for discrepancies between satellite and *in situ* values, and also inform on other locations and times where these validation statistics may be applicable. Such analyses require a significant number of match-ups, and are therefore few in number.

For the AAOT site, Mélin et al. [28] studied the dependence of validation results for $R_{rs}$ on a Case-1 versus Case-2 water partition, water single scattering albedo, angles of observation and illumination, air mass and aerosol optical thickness, using approximately 80 SeaWiFS match-

ups. The only clear dependence was found for $\tau_A$, with biases of $R_{rs}$ significantly increasing from negative to positive with increasing $\tau_A$. At the same site, for an updated atmospheric correction applied to SeaWiFS and MODIS, Zibordi et al. [29] highlighted an increase in bias and RMS difference for $R_{rs}$ in winter and for high solar zenith angles. The number of match-ups was much larger for this analysis based on field observations collected by autonomous instruments. For the same match-up data set, no significant multi-annual trends were found for $R_{rs}$ validation statistics [30]. Using the large number of match-ups found at the AAOT site, D'Alimonte et al. [31] formulated a regional model of the differences between satellite and field data of $R_{rs}$ that depended mainly on $R_{rs}$ itself; another regional model has been defined for Baltic sites [4]. This work suggests that these differences could vary according to water optical properties.

Moore et al. [32] explored this hypothesis further in the context of optical classification applied to Chla uncertainty determination. Chla uncertainty statistics were first determined for a pre-defined set of optical water types (or classes), allowing the extension of these statistics to any location on the basis of the class membership of the corresponding $R_{rs}$. Assuming that uncertainties are indeed specific to each water type, such an approach can be used to derive global maps of Chla uncertainties. Optical water types have also been used to analyze and discuss validation results for $R_{rs}$ in coastal regions [20]. A similar exercise has been performed for the match-ups gathered at the AAOT site using a set of classes defined by Moore et al. [33]. Figure 6.4 shows $R_{rs}$ associated with the considered water types (type 9 is actually an ensemble of 8 sub-types originally developed to represent coccolithophore blooms) as well as validation statistics for the optical types found in the validation data set. The RMS difference $\Delta$ increases from types 3 and 4 to type 7, while that for types 6 and 9 are close to the overall average. Relative differences tend to be lower for clearer waters (type 3) in the blue part of the spectrum (for $|\psi|$), while they are higher in the red part where the signal is lower (types 3, 4 or 6). Under the assumption that validation results obtained here are inherent to each optical water type, they could be tentatively extended to similar water types in other regions. A merit of such an approach is that its uncertainty estimates remain linked to field data.

2.4 Model-based approaches to uncertainty analysis and error propagation

Models themselves can be used to support the assessment of satellite products, a path explored mostly for bio-optical algorithms. The retrieval of bio-optical properties from ocean color radiometry often involves the spectral matching of a bio-optical model to the spectral shape of the retrieved $R_{rs}$, and this inversion process can provide valuable information on the uncertainties associated with the algorithm design or sensitivity to radiometric error. The uncertainties associated with $R_{rs}$, if known, are propagated through the bio-optical inversion and compounded by two additional factors: i) the approximations of the bio-optical model in its description of the relationship between inherent and apparent optical properties, and its

parameters describing the spectral shape of IOPs, and ii) the ambiguity of the bio-optical model, which means that the solution to the inversion is not necessarily unique [34,35].

Some studies have addressed uncertainties related to the parameters of the bio-optical model, such as the phytoplankton specific absorption, the spectral shape of the backscattering coefficient or of the absorption by chromophoric dissolved organic and detritus matter (CDM). Lee et al. [36] applied error propagation to an algebraic bio-optical model [37] to determine the uncertainties of the derived IOPs as a function of model parameters and the uncertainty of total absorption at a reference wavelength. The effect of the uncertainties associated with bio-optical parameters (e.g., defining the spectral shapes of phytoplankton and CDM absorption and backscattering coefficient) has been tested by running the inversion with different sets of parameters [38,39]. Wang et al. [38] studied the dispersion of the retrievals as a measure of output uncertainty. These approaches did not consider the other sources of uncertainties, including those of the input $R_{rs}$. Non-linear inversion of a bio-optical model provides interesting information on the uncertainty of the output IOPs from the process of minimization of a cost function [40,41,42,43]. Additionally, cases with an unsatisfactory goodness-of-fit can be filtered out as out-of-scope conditions. The uncertainty information derived from the inversion process can accompany the retrieved IOP maps and is sometimes referred to as uncertainty maps, though inversion confidence would be more appropriate terminology. The inversion confidence can account for the variance associated with the input $R_{rs}$, but it is only related to how the forward model fits the input $R_{rs}$ data, depending on the shape of the selected minimum of the cost function, and does not cope with biases affecting $R_{rs}$ or uncertainties on model formulation and parameters. Development of quality indicator maps together with derived products has also been performed using neural networks [e.g., 44].

Some atmospheric correction schemes are also based on the minimization of cost functions [45,46,47] and are amenable to the calculation of inversion confidence estimates as explored with bio-optical algorithms. Typically, these schemes have an embedded bio-optical model which constrains the distribution of retrieved $R_{rs}$. One study developed a stochastic approach to uncertainty decomposition and estimation while explicitly considering the atmospheric correction process [48]. Exercises of error propagation or accuracy analysis (e.g., [49,50]) can also provide valuable insight on atmospheric correction performance to support uncertainty assessments.

A model-based approach of a different kind has been proposed [51] making use of Chla algorithms applied to low-Chla waters. For these conditions, the assumption is that the difference in Chla computed with a standard band-ratio algorithm and with a three-band subtraction method [52] originates from uncertainties associated with $R_{rs}$. Using SeaWiFS and MODIS data, $R_{rs}$ uncertainty estimates have been expressed as a function of Chla (see Figure 6.7). Even though this approach does not apply to Chla values larger than 0.2 mg $m^{-3}$, and does not specifically account for biases, this type of technique should be further investigated.

Different methods, like those mentioned above, can inform us on various aspects of the uncertainty associated with a given retrieval, and it is desirable that their specific contributions and limitations to quantifying the overall uncertainty budget be well understood. Restricting the discussion to bio-optical algorithms producing Chla or IOPs, Table 6.1 is an attempt at broadly categorizing the type of uncertainty estimates obtained by various methods. The uncertainty on the derived product is assumed to stem from the uncertainty on the input $R_{rs}$, the potential non-uniqueness of the solution, uncertainties on the model formulation and parameters, and the uncertainties associated with the inversion process. Clearly, validation integrates all these contributions, but is affected by uncertainty in field data. Co-location techniques (see Section 3.2) share this all-encompassing character but are limited in their temporal resolution and do not consider systematic effects. Uncertainty propagation techniques can potentially accommodate uncertainties on $R_{rs}$ and model parameters, while using parameter ensembles (e.g., [38]) focuses on uncertainties on model parameters and issues of uniqueness without accounting for biases affecting $R_{rs}$. Finally, non-linear inversions provide a diagnostic of product confidence given an uncertainty on $R_{rs}$ but usually do not account for biases or parameter uncertainties. Complex approaches could combine the advantages of these various techniques.

3. Comparison of cross-mission data products

The distribution of field observations is very uneven, with a sparse coverage of the open ocean regions, particularly for optical properties. Comparison between satellite values can build upon a much larger statistical population and can support the characterization of their uncertainties. More generally, the comparison of products from different missions over their period of temporal overlap is a key element of the consistency check of the overall data record.

The comparison of two or more data products can be conducted at several levels. First, as for validation analyses, a variety of metrics may quantify the differences between a common set of data points (e.g., average difference), which is illustrated below. In the context of Earth science, it is also interesting to compare specific properties of each data record, like their spectral resolution, their spatial coverage, their inherent variability in space or time (do the data sets have the same variance, do they show the same gradients?), their seasonal cycles (do they show the same phenology?) or inter-annual signals. ~~For these properties, t~~Two data series ideally should show the same behavior for these properties, but their relative importance depends on the envisioned application. Obviously, for climate research, two data sets should provide a similar picture of seasonal to inter-annual variations.

Even if processed consistently, which means with the same principles guiding the calibration strategy and the data processing (same algorithms and binning schemes, identical ancillary data), and compared for the same day, two ocean color data sets will show differences. These result from various elements [54] including differences in sensor design and spectral characteristics as well as their implications in the specific processing codes (such as the sensitivity to polarization),

uncertainties associated with the calibration of the sensors, the sensitivity of the atmospheric correction to different aerosol types or to a different geometry of observation. Moreover, the different sensors view Earth at different times of the day, generating other sources of differences. Some might be real as associated with changes in the water properties that could be more readily studied with geostationary platforms [24]. But currently and in most cases, these differences cannot be reliably distinguished from others that are occasioned by the effect on the atmospheric correction of changes in the geometry of illumination and the atmospheric content (aerosols, clouds) or simply by noise. There is also a residual spatial mismatch as a result of the re-mapping process or of different sizes and shapes of the pixels across the satellite track. Still an additional source of differences is introduced as time composites are created since these might be built with a different temporal sampling. Finally, different products might differ because their processing chains are not consistent, e.g., with a different calibration strategy or different algorithms. With a view to create climate data records from different missions, this should be avoided as much as possible. Such a consistency is readily achieved if the different sensors share their main characteristics, for instance being similar multispectral sensors with a wavelength range of 400-900 nm as is the case for the recent global ocean color missions SeaWiFS, MODIS and MERIS. This consistency may be questioned as sensors are launched with new capabilities (channels in the ultra-violet or shortwave infrared, much higher spectral resolution, geostationary observations) that open up novel options for processing. A similar technical step took place from the Coastal Zone Color Scanner (CZCS) to the more recent sensors, prompting questions on how to process the data in a consistent manner [55].

This section reviews various approaches to compare satellite data sets. First, the issue of band shifting is addressed. Indeed, differences in center wavelengths need to be corrected for prior to comparing spectral quantities.

### 3.1 Band Shift Correction

The various ocean color missions that have been in operation have a different set of bands, which is an obstacle to a straightforward comparison of their respective records of remote sensing reflectance, $R_{rs}$. For instance the green ocean color band, which is often used as a reference band for bio-optical algorithms, is centered respectively at 547, 555 and 560 nm for the missions MODIS, SeaWiFS and MERIS, respectively. In practice it is hard to know how a MODIS $R_{rs}$ value at 547 nm compares relative to a SeaWiFS $R_{rs}$ at 555 nm. It can be done in the framework of well-defined optical properties like a Case-1 water model [56] where a certain spectral shape for $R_{rs}$ is expected for each value of the chlorophyll-a concentration (Chla) leading to the definition of a set of consistent empirical algorithms for the different sensors [57]. Such a framework could be extended to more complex optical conditions but covering the entire natural variability does not appear realistic. A similar issue arises in validation analyses when field data that are collected as multi-spectral measurements are compared with satellite $R_{rs}$. A few studies have relied on general or regional relationships to perform an action called band shift correction, whereby the $R_{rs}$ value is expressed at a target wavelength $\lambda_t$ near an existing wavelength $\lambda_0$.

The practice of band shift correction has been developed for use in validation analyses [58,4,12,15,59], comparison between satellite products [60], and as pre-processing before merging [61], with expressions linking inherent and apparent optical properties such as:

$$(7) \qquad R_{rs}(\lambda_t) = R_{rs}(\lambda_0)\frac{f(\lambda_t)}{Q(\lambda_t)}\frac{Q(\lambda_0)}{f(\lambda_0)}\frac{b_b(\lambda_t)}{a(\lambda_t)}\frac{a(\lambda_0)}{b_b(\lambda_0)}$$

where $f$ relates apparent optical properties (AOPs, irradiance reflectance) to inherent optical properties (IOPs) [62], $Q$ is the ratio of irradiance and radiance just below the surface, and $a$ and $b_b$ are the total absorption and back-scattering coefficients, respectively. Equation (**7**) requires the value of Chla to calculate $f/Q$ [63] through look-up tables computed in the framework of Case-1 water conditions.

A common requirement for these approaches is knowledge of inherent optical properties and/or concentrations of optically significant constituents sufficient to predict the spectral shape of $R_{rs}$ at least within small spectral intervals. An approach recently developed makes use of the Quasi-Analytical Algorithm (QAA) [37] to compute the absorption of phytoplankton $a_{ph}$, that of colored detrital matter (CDM) $a_{cdm}$, and the back-scattering coefficient associated with particles at 443 nm. Then IOPs are calculated at the target wavelength $\lambda_t$ using the spectral shapes of IOPs defined in the model (with the addition of the parameterization by [64] for $a_{ph}$ since QAA does not specify a spectral shape for that ~~coefficient~~property). Finally this bio-optical model is run in forward mode to calculate $R_{rs}$ at $\lambda_t$.

Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic

The results of this band shift correction are illustrated by Figure 6.5. The correction has been applied to a year (2003) of daily MODIS-Aqua $R_{rs}$ data to express them at the SeaWiFS bands. A MODIS value is computed at 510 nm by running the conversion from 488 nm and from 531 nm to 510 nm, and then taking a weighted average. All spectra common to both sensors for a given spatial bin (of a $12^{th}$-degree grid) and day have then been accumulated (49.8 million spectra). Figure 6.5a shows the overall average over that population of the SeaWiFS $R_{rs}$ and the MODIS-Aqua original $R_{rs}$ as well as the $R_{rs}$ values obtained after band shift. There is a discernible improvement in the agreement between corresponding wavelengths, at 490 and 670 nm and more clearly at 555 nm. Even the converted MODIS average value at 510 nm appears close to the SeaWiFS counterpart. Also noticeable is the fact that the use of a linear interpolation between 488 and 531 nm to compute a MODIS value at 510 nm would have resulted in a gross overestimate. Figure 6.5b is the frequency distribution of the ratio of $R_{rs}$ at the green band before and after correction, i.e., $R_{rs,A}(547)/R_{rs,S}(555)$ and $R_{rs,A}(555)/R_{rs,S}(555)$ (where A and S denote MODIS-Aqua and SeaWiFS, respectively). The median ratio decreases from 1.10 to 0.98. Band shifting is an important tool to allow inter-mission comparison, but it also contributes its own uncertainties that should be properly estimated.

3.2 Point-by-point comparison

The comparison between satellite products can be conducted for each grid point in a manner similar to validation with *in situ* data, including with the same metrics. Here again, statistics should at least provide a measure of scatter and bias, in relative terms as well as in radiometric units for $R_{rs}$. In Section 2, the quantity of reference in relative differences (i.e., the denominator) was the *in situ* value, even though *in situ* observations are not error free. In the case of a comparison between satellite products, the unbiased form of the relative difference can be preferred (in %):

$$(8) \quad |\psi^*| = 200 \frac{1}{N} \sum_{i=1}^{N} \frac{|y_i - x_i|}{x_i + y_i}$$

$$(9) \quad \psi^* = 200 \frac{1}{N} \sum_{i=1}^{N} \frac{y_i - x_i}{x_i + y_i}$$

$|\psi^*|$ and $\psi^*$ are referred to the average of the two products. The advantage is to avoid arbitrarily selecting one product as the value of reference, and numerically it prevents cases where only the denominator is close to zero. On the other hand, the difference cannot be easily interpreted in terms of a distance with respect to a clearly identified reference.

To compare SeaWiFS and MODIS products, the MODIS $R_{rs}$ data were re-binned on the SeaWiFS 12$^{th}$-degree grid, and then all daily values coincident on that grid were accumulated into 3$^{rd}$-degree macro cells. Figure 6.6a shows the resulting number of match-ups for the period 2003-2007. In general, the number of comparison data available for assessment decreases going poleward; on top of this, spatial patterns associated with persistent cloud or dust coverage are readily seen, e.g., along the inter-tropical convergence zones.

The unbiased mean relative difference $\psi^*$ is illustrated on Figure 6.6b for $R_{rs}$ at 443 nm. Most of the ocean is characterized by a relative bias not exceeding 5%, but larger differences can be noticed in specific coastal or tropical regions or in the northern Indian Ocean. Besides the spatial variations shown by inter-mission differences, comparison maps also show temporal variations, particularly changes associated with the seasonal cycle for apparent or inherent optical properties [60,65]. Examples of temporal analyses are provided in the next section.

If enough match-up data are available for a given grid point, more advanced statistics can be developed. Let us c~~C~~onsider~~ing~~ two ensembles of N coincident satellite values $(x_i)_{i=1,N}$ and $(y_i)_{i=1,N}$, each modeled as a function of a reference state *r* and zero-mean random errors δ and ε:

$$x_i = r_i + \delta_i \quad (10)$$

$$y_i = \alpha + \beta r_i + \varepsilon_i \quad (11)$$

with α and β additive and multiplicative biases, respectively, between *x* and *y*. Assuming that δ and ε are uncorrelated and independent of the reference state *r*, a mathematical development of the variance and covariance terms lead to [53]:

$$\sigma_\delta^2 = \sigma_x^2 - \frac{1}{\beta}\sigma_{xy} \quad (12)$$

$$\sigma_\varepsilon^2 = \sigma_y^2 - \beta\sigma_{xy} \quad (13)$$

which is a system of two equations with three unknowns. It can be solved with an additional assumption, for instance considering that the two satellite products, on the basis of validation analyses, have the same level of random error [53]. Solving the system may also rely on the availability of a third independent data record using a triple co-location technique. This approach is very powerful since it provides part of the uncertainty budget with the same coverage of the satellite products. Depending on the number of match-ups available, it can also be applied to separate seasons to capture variations in time.

Assuming the same level of random error for SeaWiFS and MODIS ($\sigma=\sigma_\delta=\sigma_\varepsilon$) and using the match-up data base illustrated on Figure 6.6a, a global map of σ is produced, with its global average shown on Figure 6.7. For comparison, the average over subtropical gyre waters is also given together with the uncertainty estimates for low-Chla waters given as a function of Chla by a model-based approach [51], and the unbiased RMS difference $\Delta_u$ obtained by comparison between satellite and field data at the oligotrophic MOBy site. The spectra of σ and the results obtained by the model-based approach are fairly comparable, even though the latter are higher for the case Chla=0.15 mg $m^{-3}$ in South Pacific waters. The MOBy validation results $\Delta_u$ for SeaWiFS and MODIS are also comparable with σ except in blue bands where they are closer to the model-based estimate for the case Chla=0.15 mg $m^{-3}$ in South Pacific waters. Considering the diversity of methods employed (co-location, model-based, point-wise validation), the relative agreement between these curves is interesting while the sources of differences should be further investigated.

3.3 Analysis of time series

A primary goal in development of ocean color ECVs is to enable the assessment of long-term trends to support global climate research. This leads to stringent requirements on radiometric stability, to ensure that systematic errors such as uncorrected degradation in instrument radiometric response are not misinterpreted as geophysical change. Comparative time-series analysis of $R_{rs}$ and derived products, either between satellite missions or relative to an historical reference, can identify issues with instrument radiometric characterization and temporal

calibration stability. Analysis of the seasonal trends observed in different latitudinal zones for MODIS-Aqua $R_{rs}$ time-series relative to SeaWiFS, for example, contributed to the discovery of an error in characterization of polarization sensitivity on MODIS [66]. Without the SeaWiFS time-series for comparison, this error may have never been identified, and seasonal cycles in the ocean color signal from MODIS in climate critical high-latitude regions would have been highly misleading. For products derived using common algorithms, relative agreement between missions also provides a measure of uncertainty for trend detection. Franz et al. [67], for example, used the average difference in regional monthly means between consistently processed SeaWiFS and MODIS data as a measure of uncertainty in the 15-year multi-mission time-series of Chla.

A typical time-series analysis starts with the data product of interest projected into a set of fixed geographic bins and averaged over specific temporal intervals. A widely used example is the SeaWiFS 9.2-km binned product: a globally-distributed set of quasi-equal-area bins where the value of each bin represents the local product average over 8-day or monthly time intervals [68]. The global dataset or a subset of the bins (e.g., based on geography or water-type classification) is then spatially averaged within each time interval, and the averages trended in time. The preferred time interval for compositing is a trade-off between minimizing the geophysical variability lost to the average and maximizing the number of observed (or filled) bins. When comparing time-series between missions, it is also useful to first reduce the selected bins within each time-interval to a set of common filled bins. This is critical for the identification of anomalous sensor-calibration artifacts, as some missions show systematic geographic gaps even after 8-days of compositing, and these geographic sampling biases induce additional variability in mission-to-mission differences.

As an example, Figure 6.8 shows $R_{rs}$ trends from MODIS on Terra and MODIS on Aqua, based on common bins over the overlapping missions. The measurements were restricted to include only those bins where water-depth is greater than 1000 meters (to avoid the complexities and diurnal variability of the coastal regions), and an average was computed over this geographic subset for each month of the time-series. The comparison clearly demonstrates a degradation in the radiometric stability of MODIS-Terra relative to MODIS-Aqua that was traced to the MODIS-Terra instrument calibration and subsequently corrected.

For a radiometrically stable sensor, the dominant variability in the derived $R_{rs}$ time-series for the deep oceans is a seasonal cycle associated with phytoplankton productivity. Subtraction of this mean seasonal cycle from the $R_{rs}$ trends yields an anomaly time-series. While anomaly trends provide a mechanism for investigating long-term geophysical changes in the ocean color record, they can also serve as a powerful tool to identify sensor radiometric instabilities. Figure 6.9a shows the anomaly in MODIS-Aqua $R_{rs}(547)$ relative to the mean seasonal cycle for the deep ocean gyres. At this wavelength, the $R_{rs}$ signal is relatively insensitive to small changes in Chla, and so we expect the time-series in these very low productivity regions to show little variability,

as is the case for MODIS. For the 560-nm band of MERIS, however, the $R_{rs}$ anomaly time-series shows a strong deviation in 2005-2006 suggesting a 5-10% bias that was traced to a change in the operating state of the instrument.

3.4 Climate signal analysis

Ocean color products are being scrutinized across a whole range of space and time scales. The various satellite data records should show the same patterns of variability, seasonal cycle (phenology), and trends. The global distribution of phytoplankton is well known and reproduced by all satellite products, but the advent of high resolution modeling and remote sensing is shedding new light on how phytoplankton and physics are related across spatial scales, including planetary waves [69], mesoscale and submesoscale [70,71,72], or internal waves [73]. The seasonal cycle of phytoplankton is the most prominent signal in many ocean regions and phytoplankton phenology has recently been actively investigated [e.g., 74]. Variations in Chla distributions have been studied at other time scales, describing intra-seasonal signals such the Madden-Julian oscillation [75], tropical inter-annual variations like El Niño [76] or climate signals with longer time scales [77]. How the different characteristics of each satellite product, in terms of spatial resolution, levels and structure of variance, or noise, affect analyses of climate signals or model simulations relying on data assimilation has been largely unexplored. These differences will have to be properly integrated into the long-term analysis of the biogeochemical responses of marine ecosystems to climate forcing.

However, some studies have analyzed how different satellite missions represent the temporal evolution of a satellite derived product (Chla or optical properties) for specific regions [e.g., 78,79,65]. Djavidnia et al. [80] have compared Chla time series averaged over the Longhurst [81] provinces, as obtained from SeaWiFS, MODIS and MERIS. Taylor plots are useful in that regard, illustrating on the same plot the correlation between two signals, their standard deviations, and their unbiased RMS difference. From the updated results for SeaWiFS and MODIS seen on Figure 6.10, it appears that the correlation coefficients between monthly series are all higher than 0.8, while the variance of the MODIS time series can be lower or higher than that of SeaWiFS.

As soon as the SeaWiFS mission lifetime exceeded 5 years, investigations started studying possible trends associated with its Chla series [e.g., 82,83]. The validity of these analyses was supported by the activities that ensured the characterization and stability of the instrument calibration [84]. Even if similar calibration strategies are followed for the main ocean color missions, it still appears worthwhile to compare long-term trends obtained from different missions to check that they provide the same view of inter-annual changes taking place in the oceans. Generally, this is unlikely to be an easy task since it requires overlaps between missions long enough for trend analysis. But the ocean color community has been fortunate to benefit from such cases with the long records of SeaWiFS, MODIS and MERIS. For instance, the latter two missions were contemporaneously in operation for a decade. A trend analysis was performed

on the MODIS and MERIS Chla data over the period August 2002 to July 2011 using a non-parametric seasonal Kendall test [77]. Figure 6.11 illustrates the agreement between the trend fields. The use of a contingency matrix allows the quantification of this agreement by computing the percentage of the ocean with similar or divergent behaviors (trend slopes of the same/opposite signs, significance levels, …). For instance, 20% of the ocean is found to have a statistically significant trend ($p<0.05$) of the same sign for both series (11% with a positive slope, 9% with a negative slope), while there is virtually no area with a statistically significant trend of opposite sign (0.005%).

Analyses checking the consistency of ocean color time series in terms of spatial distributions, phenology or trends should be seen as integral parts of an assessment strategy applied to climate data records.

## 4. Conclusions

In the assessment of satellite products, a recurrent question is that of ranking: is a given product better than another, either because processed with different algorithms or associated with another sensor? A still larger question is: should it be deemed acceptable for climate research? Validation statistics might promote one mission for one specific product (e.g., Chla but not IOPs), using one field data set but not another, for some wavelengths only or for the spectral shape of $R_{rs}$. One product might also be preferred for its extensive data coverage. Ultimately, the choice of a particular product is intimately linked with the science question being addressed, and its assessment needs to be commensurate with the intended application. The stringent requirements associated with climate research call for a comprehensive approach including validation with field data, uncertainty analyses, and consistency checks like comparison between mission-specific products and time series analyses. In that regard, significant mission overlaps (at least one year) are an absolute prerequisite for such a strategy, besides the fact that gaps in the data records would seriously challenge our ability to use the ocean color record to detect climate signals [85].

In 1983, Gordon et al. [86] could use three spectra of water leaving radiance determined from ship-based observations to assess the atmospheric correction applied to CZCS data. Although this chapter illustrates the considerable progress made in data collection since then, the relative scarcity of high-quality $R_{rs}$ measurements remains a limiting factor for assessing radiometric satellite ocean color products. The oceanographic community should invest in comprehensive measurement programs for validation purposes, and in the development of new technological or methodological approaches. The development of a network of automated above-water radiometers [13] represents a major progress for coastal waters. Placing bio-optical instruments onto floats [87] is also a promising avenue to increase frequency and coverage. Hyperspectral measurements are also desirable, both to fully accommodate the spectral characteristics specific to each satellite sensor, and in preparation for advanced spaceborne sensors with hyperspectral capabilities.

This chapter has also discussed approaches based on the use of models or inter-comparison techniques that can profitably complement the *in situ* validation statistics. Their major contribution should be to allow an extension of validation results to a wider range of geometric and environmental conditions. A complete framework for error propagation is needed, but it requires a thorough and accurate characterization of the uncertainties associated with the radiance signal at top-of-atmosphere, as well as a detailed understanding of the propagation of errors through the atmospheric correction algorithm, the uncertainties associated with algorithm assumptions, and the contributions of all other ancillary inputs (e.g., meteorological conditions and atmospheric gas contributions) to the total uncertainty budget. The approaches mentioned here do not form an exhaustive list. More holistic modeling environments could be devised to support product assessment, for instance to exclude some conditions or ascertain their probability. For instance, high Chla concentrations close to river outlets are unlikely in winter, and their presence in satellite products is suspect. Ecosystem models can contribute to assess and improve satellite products, and vice-versa.

A maturity model has been proposed to assess the completeness of climate data records (CDR) [88]. This matrix model contemplates six levels of maturity for all aspects of a data set. We could argue that the ocean color products can pretend to the levels 5 or 6, synonymous of a “full operational capability”, for aspects like software, documentation, metadata, public access and utility. Understandably, product validation is lagging behind, qualifying for levels 3 or 4, “uncertainty estimated for select locations/times” or “uncertainty estimated over widely distributed times/locations by multiple investigators; differences understood”. Level 5 would entail the knowledge of “consistent uncertainties estimated over most environmental conditions by multiple investigators”. Progress has been made in that direction but needs to be consolidated into an operational context to allow a fully informed use of ocean color products in climate research.

Acknowledgments

The authors would like to thank all the investigators contributing to the collection of in-situ data used in this chapter, particularly for the AERONET-OC and MOBy sites. Their dedication is earnestly valued.

References

[1] GCOS-154, “Systematic observation requirements for satellite-based products for climate,” supplemental details to the satellite-based component of the “Implementation plan for the Global Observing System for Climate in Support of the UNFCC, 138 pp. (2011).
[2] IOCCG, “Remote sensing of optical properties: Fundamentals, tests of algorithms, and applications,” Z.-P. Lee (Ed), Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada, 126 pp. (2006).
[3] S.W. Bailey, and P.J. Werdell, “A multi-sensor approach for the on-orbit validation of ocean color satellite data products,” Remote Sens. Environ. **102**, 12-23 (2006).
[4] G. Zibordi, J.-F. Berthon, F. Mélin, D. D'Alimonte, and S. Kaitala, “Validation of satellite ocean color primary products at optically complex coastal sites: northern Adriatic Sea, northern Baltic Proper, Gulf of Finland,” Remote Sens. Environ. **113**, 2574-2591 (2009).
[5] H. Feng, D. Vandemark, J.W. Campbell, B.N. Holben, “Evaluation of MODIS ocean colour products at a northeast United States coast site near the Martha's Vineyard Coastal Observatory,” Int. J. Remote Sens. **29**, 4479-4497 (2009).
[6] R. Brewin, S. Sathyendranath, D. Mueller, C. Brockmann, P.-Y. Deschamps, E. Devred, R. Doerffer, N. Fomferra, B. Franz, M. Grant, S. Groom, A. Horseman, C. Hu, H. Krasemann, Z.-P. Lee, S. Maritorena, F. Mélin, M. Peters, T. Platt, P. Regner, T. Smyth, F. Steinmetz, J. Swinton, P.J. Werdell, and G.N. White, “The Ocean Colour Climate Change Initiative. III. A round-robin comparison on in-water bio-optical algorithms,” Remote Sens. Environ., 10.1016/j.rse.2013.09.016 (2014).
[7] D. Mueller, H. Krasemann, R. Brewin, C. Brockmann, P.-Y. Deschamps, R. Doerffer, N. Fomferra, B.A. Franz, M. Grant, S. Groom, F. Mélin, T. Platt, P. Regner, S. Sathyendranath, and F. Steinmetz, “The Ocean Colour Climate Change Initiative. II An assessment of atmospheric correction processors based on in-situ measurements,” Remote Sens. Environ., in press (2014).
[8] D.K. Clark, H.R. Gordon, K.J. Voss, Y. Ge, W. Broenkow, and C. Trees, “Validation of atmospheric correction over the oceans,” J. Geophys. Res. **102**, 17209-17217 (1997).
[9] B.A. Franz, S.W. Bailey, P.J. Werdell, and C.R. McClain, “Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry,” Appl. Opt, **46**, 5068-5082 (2007).
[10] W.W. Gregg, and N.W. Casey, “Global and regional evaluation of the SeaWiFS chlorophyll data set,” Remote Sens. Environ. **93**, 463-479 (2004).
[11] P.J. Werdell, S.W. Bailey, G. Fargion, C. Pietras, K. Knobelspiesse, G.C. Feldman, and C.R. McClain, “Unique data repository facilitates ocean color satellite validation,” EOS Trans. AGU **84**, 379 (2003).
[12] G. Zibordi, J.-F. Berthon, F. Mélin, and D. D’Alimonte, “Cross-site consistent in-situ measurements for satellite ocean color applications: the BiOMaP radiometric dataset. Remote Sens. Environ. **115**, 2104-2115 (2011).
[13] G. Zibordi, B.N. Holben, S.B. Hooker, F. Mélin, J.-F. Berthon, I. Slutsker, D. Giles, D. Vandemark, H. Feng, K. Rutledge, G. Schuster, A. Al Mandoos, “A network for standardized ocean color validation measurements,” EOS Trans. AGU **87**, 293, 297 (2006).
[14] T. Cui, J. Zhang, S. Groom, L. Sun, T. Smyth, and S. Sathyendranath, “Validation of MERIS ocean-color products in the Bohai Sea: A case study for coastal waters,” Remote Sens. Environ. **114**, 2326-2336 (2010).
[15] D. Antoine, F. d'Ortenzio, S.B. Hooker, G. Bécu, B. Gentili, D. Tailliez, and A.J. Scott, “Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color

sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Sea (BOUSSOLE project). J. Geophys. Res. **113**, C07013, 10.1029/2007JC004472 (2008).
[16] M.E. Smith, S. Bernard, and S. O'Donoghue, "The assessment of optimal MERIS ocean colour products in the shelf waters of the KwaZulu-Natal Bight, South Africa," Remote Sens. Environ. **137**, 124-138 (2013).
[17] G. Zibordi, F. Mélin, J.-F. Berthon, and E. Canuti, "Assessment of MERIS ocean color data products for European seas," Ocean Sci. **9**, 521-533 (2013).
[18] H. Murakami, K. Sasaoka, K. Hosoda, H. Fukushima, M. Toratani, R. Frouin, B.G. Mitchell, M. Kahru, P.-Y. Deschamps, D. Clark, S. Flora, M. Kishino, S.-I. Saitoh, I. Asanuma, A. Tanaka, H. Sasaki, K. Yokouchi, Y. Kiyomoto, H. Saito, C. Dupouy, A. Siripong, S. Matsumura, and H. Ishizaka, "Validation of ADEOS-II GLI ocean color products using in-situ observations," J. Oceanogr. **62**, 373-393 (2006).
[19] C. Jamet, H. Loisel, C.P. Kuchinke, K. Ruddick, G. Zibordi, and H. Feng, "Comparison of three SeaWiFS atmospheric correction algorithms for turbid waters using AERONET-OC measurements," Remote Sens. Environ. **115**, 1955-1965 (2011).
[20] C. Goyens, C. Jamet, and T. Schroeder, "Evaluation of four atmospheric correction algorithms for MODIS-Aqua images over contrasted coastal waters", Remote Sens. Environ. **131**, 63-75 (2013).
[21] M. Shimada, H. Oaku, Y. Mitomi, H. Murakami, A. Mukaida, Y. Nakamura, J. Ishisaka, H. Kawamura, T. Tanaka, M. Kishino, and H. Fukushima, "Calibration and validation of the ocean color version-3 product from AEOS OCTS," J. Oceanogr. **54**, 401-416 (1998).
[22] S. Hlaing, T. Harmel, A. Gilerson, R. Foster, A. Weidemann, R. Arnone, M. Wang and S. Ahmed, "Evaluation of the VIIRS ocean color monitoring performance in coastal regions," Remote Sens. Environ. **139**, 398-414 (2013).
[23] N. Lamquin, C. Mazeran, D. Doxaran, J.-H. Ryu, and Y.-J. Park, "Assessment of GOCI radiometric products using MERIS, MODIS and field measurements," Ocean Sci. J. **47**, 287-311 (2012).
[24] M. Wang, J.H. Ahn, L. Jiang, W. Shi, S. Son, Y.-J. Park, and J.-H. Ryu, "Ocean color products from the Korean Geostationary Ocean Color Imager (GOCI)," Opt. Exp **21**, 3835-3849 (2013).
[25] T. Cui, J. Zhang, J. Tang, S. Sathyendranath, S. Groom, Y. Ma, W. Zhao, and Q. Song, "Assessment of satellite ocean color products of MERIS, MODIS and SeaWiFS along the East China Coast (in the Yellow Sea and East China Sea)," ISPRS J. Photogram. Remote Sens. **87**, 137-151 (2014).
[26] T. Schroeder, I. Behnert, M. Schaale, J. Fischer, and R. Doerffer, "Atmospheric correction algorithm for MERIS above case-2 waters," Int. J. Remote Sens. 28, 1469-1486 (2007).
[27] M. Wang, S.-H. Son, and W. Shi, "Evaluation of MODIS SWIR and NIR-SWIR atmospheric correction algorithms using SeaBASS data," Remote Sens. Environ. **113**, 635-644 (2009).
[28] F. Mélin, G. Zibordi, and J.-F. Berthon, "Assessment of satellite ocean color products at a coastal site," Remote Sens. Environ. **110**, 192-215 (2007).
[29] G. Zibordi, F. Mélin, and J.-F. Berthon, "Intra-annual variations of biases in remote sensing primary ocean color products at a coastal site," Remote Sens. Environ **124**, 627-636 (2012).
[30] G. Zibordi, F. Mélin, and J.-F. Berthon, "Trends in the bias of primary satellite ocean color products at a coastal site," IEEE Geosci. Remote Sens. Lett. **9**, 1056-1060 (2012).

[31] D. D'Alimonte, G. Zibordi, and F. Mélin, "A statistical method for generating cross-mission consistent normalized water-leaving radiances," IEEE Trans. Geosci. Remote Sens. **46**, 4075-4093 (2008).
[32] T.S. Moore, J.W. Campbell, and M.D. Dowell, "A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product," Remote Sens. Environ. **113**, 2424-2430 (2009).
[33] T.S. Moore, M.D. Dowell, and B.A. Franz, "Detection of coccolithophore blooms in ocean color satellite imagery: A generalized approach for use with multiple sensors," Remote Sens. Environ. **117**, 249-263 (2012).
[34] M. Sydor, R.W. Gould, R.A. Arnone, V.I. Haltrin, and W. Goode, "Uniqueness in remote sensing of the inherent optical properties of ocean water," Appl. Opt. **43**, 2156-2162 (2004).
[35] M. Defoin-Platel, and M. Chami, "How ambiguous is the inverse problem of ocean color on coastal waters?" J. Geophys. Res. **112**, C03004, 10.1029/2006JC003847 (2007).
[36] Z.-P. Lee, R.A. Arnone, C. Hu, P.J. Werdell, and B. Lubac, "Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm. Appl. Opt. **49**, 369-381 (2010).
[37] Z.-P. Lee, K.L. Carder, and R.A. Arnone, "Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters," Appl. Opt. **41**, 5755-5772 (2002).
[38] P. Wang, E.S. Boss, and C. Roesler, "Uncertainties of inherent optical properties obtained from semi-analytical inversions of ocean color," Appl. Opt. **44**, 4074-4085 (2005).
[39] V. Brando, A.G. Dekker, Y.J. Park, and T. Schroeder, "Adaptive semi-analytical inversion of ocean color radiometry in optically complex waters," Appl. Opt. **51**, 2808-2833 (2012).
[40] H.J. Van Der Woerd, and R. Pasterkamp, "HYDROPT: A fast and flexible method to retrieve chlorophyll-a from multispectral satellite observations of optically complex coastal waters," Remote Sens. Environ. **112**, 1795-1807 (2008).
[41] M.S. Salama, A.G. Dekker, Z. Su, C.M. Mannaert,s and W. Verhoef, "Deriving inherent optical properties and associated inversion-uncertainties in the Dutch lakes," Hydrol. Earth Syst. Sci. **13**, 1113-1121 (2009).
[42] S. Maritorena, O. Hembise Fanton d'Andon, A. Mangin, and D.A. Siegel, "Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues," Remote Sens. Environ. **114**, 1791-1804 (2010).
[43] P.J. Werdell, B.A. Franz, S.W. Bailey, G.C. Feldman, E. Boss, V.E. Brando, M.D. Dowell, T. Hirata, S.J. Lavender, Z.-P. Lee, H. Loisel, S. Maritorena, F. Mélin, T.S. Moore, T.J. Smyth, D. Antoine, E. Devred, O. Fanton d'Andon, and A. Mangin, "A generalized ocean color inversion for retrieving marine inherent optical properties," Appl. Opt. **52**, 2019-2037 (2013).
[44] H. Schiller, and R. Doerffer, "Improved determination of coastal water constituent concentrations from MERIS data," IEEE Trans. Geosci. Remote Sens. **43**, 1585-1591 (2005).
[45] R.M. Chomko, and H.R. Gordon, "Atmospheric correction of ocean color imagery: Test of the spectral optimization algorithm with the Sea-viewing Wide Field-of-View Sensor," Appl. Opt. **40**, 2973-2984, 2001.
[46] K. Stamnes, W. Li, B. Yan, H. Eide, A. Barnard, W.S. Pegau, and J.J. Stamnes, "Accurate and self-consistent ocean color algorithms: Simultaneous retrieval of aerosol optical properties and chlorophyll concentrations," Appl. Opt. **42**, 939-951 (2003).
[47] C. Jamet, S. Thiria, C. Moulin, and M. Crépon, "Use of a neurovariational inversion for retrieving oceanic and atmospheric constituents from ocean color imagery: A feasibility study,"

J. Atmos. Ocean. Tech. **22**, 460-475 (2005).
[48] M.S. Salama, and A. Stein, "Error decomposition and estimation of inherent optical properties," Appl. Opt. **48**, 4947-4962 (2009).
[49] B. Bulgarelli, and G. Zibordi, "Remote sensing of ocean colour: accuracy assessment of an approximate atmospheric correction code," Int. J. Remote Sen. **24**, 491-509 (2003).
[50] B. Bulgarelli, F. Mélin, and G. Zibordi, "SeaWiFS-derived products in the Baltic Sea: performance analysis of a simple atmospheric correction algorithm," Oceanologia **45**, 655-677 (2003).
[51] C. Hu, L. Feng, and Z.-P. Lee, "Uncertainties of SeaWiFS and MODIS remote sensing reflectance: Implications from clear water assessments," Remote Sens. Environ. **133**, 168-182 (2013).
[52] C. Hu, Z.-P. Lee, and B.A. Franz, "Chlorophyll *a* algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference," J. Geophys. Res. **117**, C01011, 10.1029/2011JC007395 (2012).
[53] F. Mélin, "Global distribution of the random uncertainty associated with satellite-derived Chla," IEEE Geosci. Remote Sens. **7**, 220-224 (2010).
[54] IOCCG, "Ocean Colour Data Merging," W.W. Gregg (Ed), J. Aiken, E. Kwiatkowska, S. Maritorena, F. Mélin, H. Murakami, S. Pinnock, C. Pottier, Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada, 65pp. (2007).
[55] D. Antoine, A. Morel, H.R. Gordon, V.F. Banzon, and R.H. Evans, "Bridging ocean color observations of the 1980s and 2000s in search of long-term trends," J. Geophys. Res. **110**, C06009, 10.1029/2004JC002620 (2005).
[56] A. Morel, and S. Maritorena, "Bio-optical properties of oceanic waters: A reappraisal," J. Geophys. Res. **106**, 7163-7180 (2001).
[57] A. Morel, Y. Huot, B. Gentili, P.J. Werdell, S.B. Hooker, and B.A. Franz, "Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach," Remote Sens. Environ. **111**, 69-88 (2007).
[58] G. Zibordi, F. Mélin, and J.-F. Berthon, "Comparison of SeaWiFS, MODIS and MERIS radiometric products at a coastal site," Geophys. Res. Lett. **33**, L06617, 10.1029/2006GL025778, (2006).
[59] F. Mélin, G. Zibordi, J.-F. Berthon, S.W. Bailey, B.A. Franz, K.J. Voss, S. Flora, and M. Grant, "Assessment of MERIS reflectance data as processed by SeaDAS over the European Seas," Opt. Exp. **19**, 25657-25671 (2011).
[60] F. Mélin, G. Zibordi, and S. Djavidnia, "Merged series of normalized water leaving radiances obtained from multiple satellite missions for the Mediterranean Sea," Adv. Space Res. **43**, 423-437 (2009).
[61] F.Mélin, V. Vantrepotte, M. Clerici, D. D'Alimonte, G. Zibordi, J.-F. Berthon, and E. Canuti, "Multi-sensor satellite time series of optical properties and chlorophyll *a* concentration in the Adriatic Sea," Prog. Oceanogr. **91**, 229-244 (2011).
[62] A. Morel, and B. Gentili, "Diffuse reflectance of oceanic waters: its dependence on Sun angle as influenced by the molecular scattering contribution," Appl. Opt. **30**, 4427-4438 (1991).
[63] A. Morel, D. Antoine, and B. Gentili, "Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function," Appl. Opt. **41**, 6289-6306, (2002).

[64] A. Bricaud, M. Babin, A. Morel, and H. Claustre, "Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization," J. Geophys. Res. **100**, 13321-13332 (1995).
[65] F. Mélin, "Comparison of SeaWiFS and MODIS time series of inherent optical properties for the Adriatic Sea," Ocean Sci. **7**, 351-361 (2011).
[66] G. Meister, E.J. Kwiatkowska, B.A. Franz, F.S. Patt, G.C. Feldman, and C.R. McClain, "Moderate Resolution Imaging Spectroradiometer ocean color polarization correction," Appl. Opt. **44**, 5524-5535 (2005).
[67] B.A. Franz, D.A. Siegel, M.J. Behrenfeld, and P.J. Werdell, "Global ocean phytoplankton," in State of the Climate 2012, Bull. Am. Meteorol. Soc. **94**, S75-S78 (2013).
[68] J.W. Campbell, J.M. Blaisdell, and M. Darzi, "Level-3 SeaWiFS data products: Spatial and temporal binning algorithms," NASA Tech. Mem. 104566, vol. 32, Eds. S.B. Hooker, E.R. Firestone and J.G. Acker, NASA Goddard Space Flight Center, Greenbelt, Maryland (1995).
[69] P. Cipollini, D. Cromwell, P.G. Challenor, and S. Raffaglio, "Rossby waves detected in global ocean colour data," Geophys. Res. Lett. **28**, 323-326 (2001).
[70] A. Mahadevan, and J.W. Campbell, "Biogeochemical patchiness at the sea surface," Geophys. Res. Lett. **29**, 1926, 10.1029/2001GL014116 (2002).
[71] S.C. Doney, D.M. Glover, S.J. McCue, and M. Fuentes, "Mesoscale variability of Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite ocean color: Global patterns and spatial scales," J. Geophys. Res. **108**, 3024, 10.1029/2001JC000843 (2003).
[72] M. Lévy, R. Ferrari, P.J.S. Franks, A.P. Martin, P. Rivière, "Bringing physics to life at the submesoscale," Geophys Res. Lett. **39**, L14062, 10.1029/2012GL052756 (2012).
[73] J.C.B. da Silva, A.L. New, M.A. Srokosz, and T.J. Smyth, "On the observability of internal tidal waves in remotely-sensed ocean colour data," Geophys. Lett. 29, 1569, 10.1029/2001GL013888 (2002).
[74] M.R.P. Sapiano, C.W. Brown, S. Schollaert Uz, and M. Vargas, "Establishing a global climatology of marine phytoplankton phenological characteristics. J. Geophys. Res. **117**, C08026, 10.1029/2012JC007958 (2012).
[75] D. Jin, D.E. Waliser, C. Jones, and R. Murtugudde, "Modulation of tropical ocean surface chlorophyll by the Madden-Julian Oscillation," Clim. Dyn. **40**, 39-58 (2013).
[76] M.J. Behrenfeld, J.T. Randerson, C.R. McClain, G.C. Feldman, S.O. Los, C.J. Tucker, P.G. Falkowski, C.B. Field, R. Frouin, W.E. Esaias, D.D. Kolber, N.H. Pollack, "Biospheric primary production during an ENSO transition. Science **291**, 2594-2597 (2001).
[77] V. Vantrepotte, and F. Mélin, "Inter-annual variations in the SeaWiFS global chlorophyll *a* concentration (1997-2007)," Deep-Sea Res. I **58**, 429-441 (2011).
[78] C. Zhang, C. Hu, S. Shang, F.E. Mueller-Karger, Y. Li, M. Dai, B. Huang, X. Ning, and H. Hong, H., "Bridging between SeaWiFS and MODIS for continuity of chlorophyll-a concentration assessments off Southeastern China. Remote Sens. Environ. **102**, 250-263 (2006).
[79] P.J. Werdell, S.W. Bailey, B.A. Franz, L.W. Harding, G.C. Feldman, and C.R. McClain, "Regional and seasonal variability of chlorophyll-a in Chesapeake Bay as observed by SeaWiFS and MODIS-Aqua," Remote Sens. Environ. **113**, 1319-1230 (2009).
[80] S. Djavidnia, F. Mélin, and N. Hoepffner, "Comparison of global ocean colour data records," Ocean Sci. **6**, 61-76 (2010).
[81] A. Longhurst, "Ecological geography of the sea," Academic Press, 560 pp. (2006).
[82] C.R. McClain, S.R. Signorini, and J.R. Christian, "Subtropical gyre variability observed by ocean-color satellites," Deep-Sea Res. II **51**, 281-301 (2004).

[83] W.W. Gregg, N.W. Casey, and C.R. McClain, “Recent trends in global ocean chlorophyll,” Geophys Res. Lett. **32**, L03606, 10.1029/2004GL021808 (2005).
[84] R.E. Eplee, G. Meister, F.S. Patt, R.A. Barnes, S.W Bailey, B.A. Franz, and C.R. McClain, “On-orbit calibration of SeaWiFS,”, Appl. Opt. **51**, 8702-8730 (2012).
[85] C. Beaulieu, S.A. Henson, J.L. Sarmiento, J.P. Dunne, R.R. Rykaczewski, and L. Bopp, “Factors challenging our ability to detect long-term trends in ocean chlorophyll” Biogeosciences **10**, 2711-2724 (2013).
[86] H.R. Gordon, D.K. Clark, J.W. Brown, O.B. Brown, R.H. Evans, and W.W. Broenkow, “Phytoplankton pigment concentrations in the Middle Atlantic Bight: comparison between ship determinations and Coastal Zone Color Scanner estimates,” Appl. Opt. **22**, 20-36 (1983).
[87] IOCCG “Bio-optical sensors on Argo floats,” H. Claustre (Ed), Reports of the International Ocean-Colour Coordinating Group, No. 11, IOCCG, Dartmouth, Canada, 89 pp. (2011).
[88] J.J. Bates, and J.L Privette, “A maturity model for assessing the completeness of climate data records”, EOS Trans. Am. Geophys. Union **93**, 441 (2012).

Tables

| | $R_{rs}$ | Uniqueness | Parameters | Inversion |
|---|---|---|---|---|
| Validation | x,t | | | |
| Uncertainty propagation [36] | X,T | | X,T | |
| Parameters Ensemble [38] | | X,T | | |
| Non-Linear Inversion [42] | X,T | | | X,T |
| Co-location [53] | X,t | | | |

Table 6.1: Matrix relating error sources affecting products of bio-optical algorithms and different methods computing uncertainty terms. Small letters refer to results obtained at selected locations x and times t, while capital letters indicate estimates potentially obtained at each pixel. Blue color indicates that the approach can handle systematic effects (biases). References are only intended as general examples for a given approach.

Figures

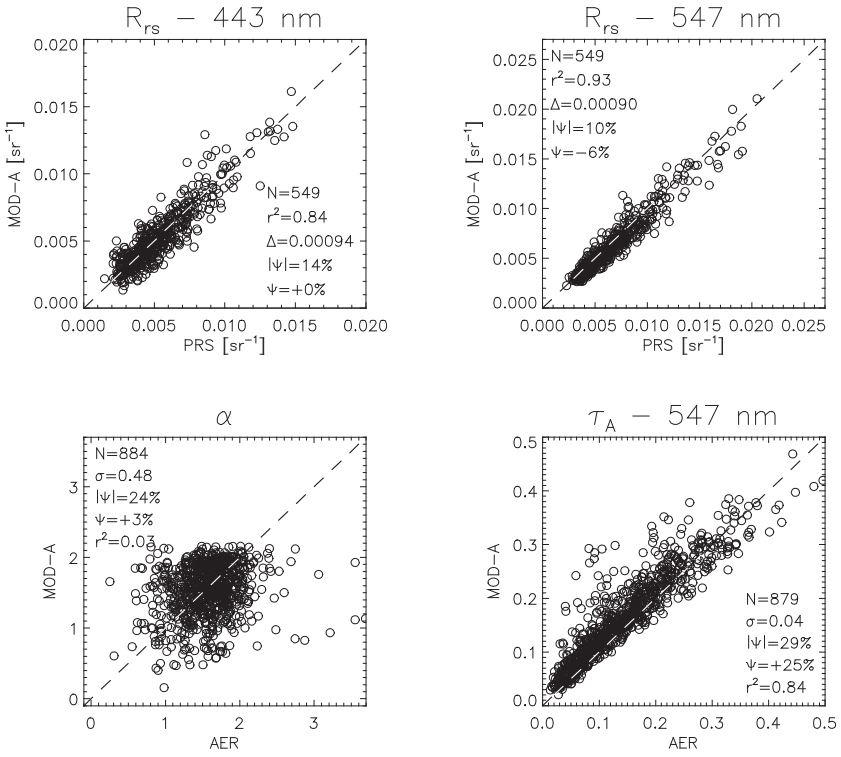


Figure 6.1: Comparison between above-water radiometry and MODIS-Aqua products for $R_{rs}$ at 443 and 547 nm, and aerosol optical thickness $\tau_A$ at 547 nm and Ångström exponent α. Validation statistics are introduced in Section 2.2.

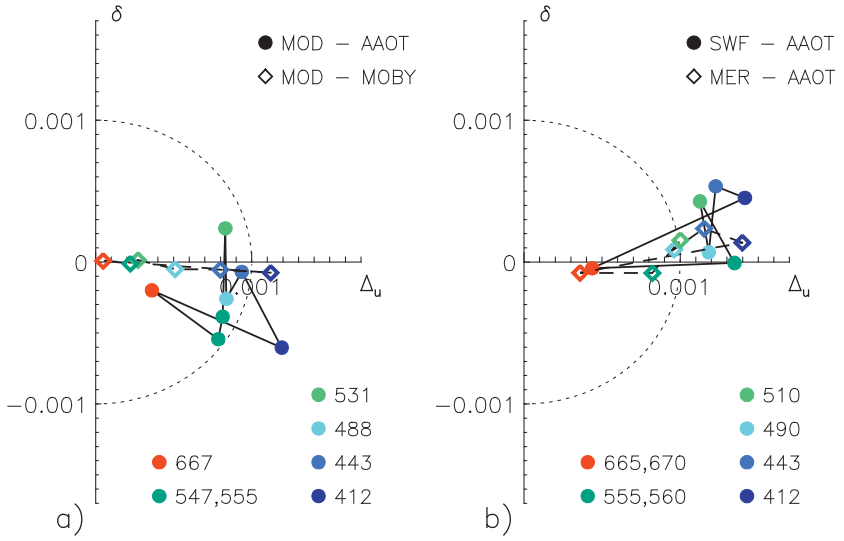


Figure 6.2: Spectral target diagram for validation results for MODIS-Aqua a) at MOBy (N=229) and AAOT (N=549, except at 531 nm, N=176) and b) SeaWiFS (N=369) and MERIS (N=149) at AAOT (right). Axes are in units of $sr^{-1}$. See text for definition of statistical quantities.

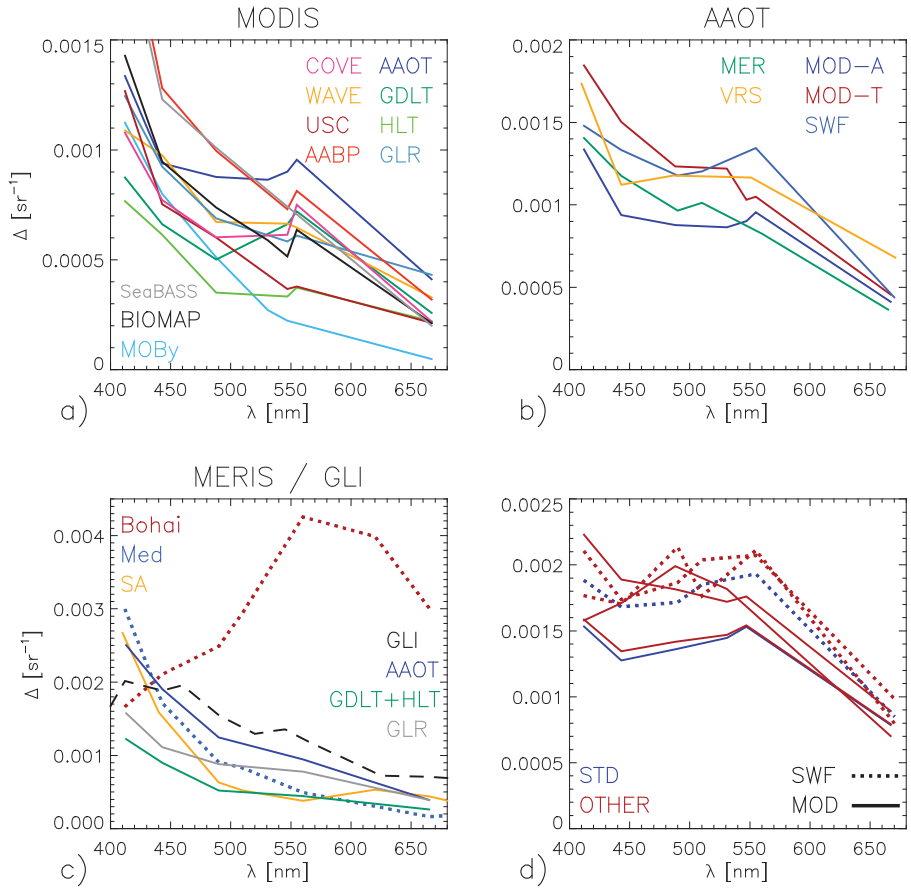


Figure 6.3: RMS differences between satellite and field data of $R_{rs}$ (in $sr^{-1}$): a) MODIS-Aqua compared to various data sets from AERONET-OC sites (see text) as well as SeaBASS, BiOMaP and MOBy; b) Results for MODIS-Aqua (N=549) and Terra (270), MERIS (N=149), SeaWiFS N=369) and VIIRS (N=70) at AAOT; c) Results for MERIS processed by ESA processor MEGS (version 7 as dotted line, 8 otherwise) given for the Bohai Sea [14], the northwest Mediterranean Sea [15], South African (SA) coastal waters [16], and AERONET-OC sites, AAOT (N=86), GDLT and HLT (Baltic Sea) and GLR (Black Sea) [17]. GLI results are represented by the black dashed line [18]. d) Results obtained in coastal waters for SeaWiFS (dotted line) [19] and MODIS-Aqua [20] with different atmospheric correction schemes, including the standard SeaDAS (STD, in blue). When appropriate the source reference is given.

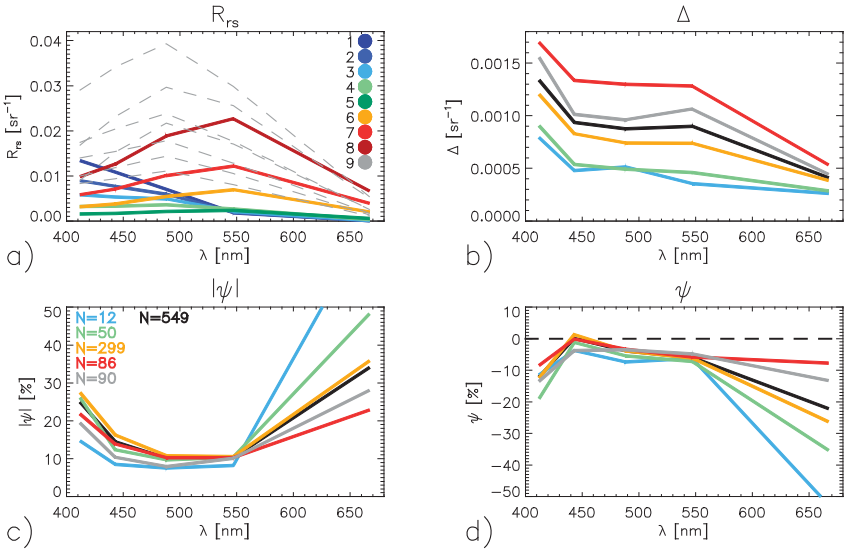


Figure 6.4: Dependence of validation statistics on optical water types at AAOT. a) Mean $R_{rs}$ for the optical water types defined in [33]; type 9 includes 8 sub-types. Validation results computed by types for b) RMS difference Δ, c) mean absolute relative difference |ψ|, and d) mean relative difference ψ. Results for all match-ups combined are in black; results per optical type are shown with the same color code as for a) only if the members number at least 10 spectra.

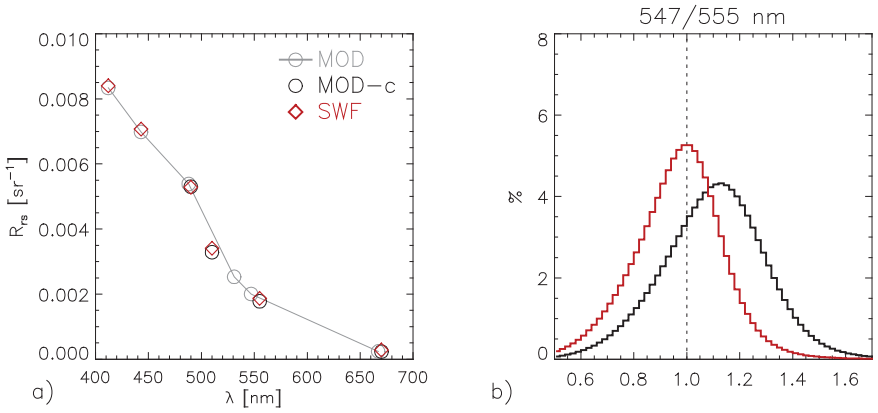


Figure 6.5: a) Average of all daily $R_{rs}$ coincident between MODIS-Aqua and SeaWiFS in 2003. Curve with grey circles represent MODIS values without band shifting, while black circles represent MODIS statistics computed after the band shifting correction has been applied. b) Histogram of the ratio between MODIS (with and without band shifting, in red and black, respectively) and SeaWiFS $R_{rs}$ in the green band.

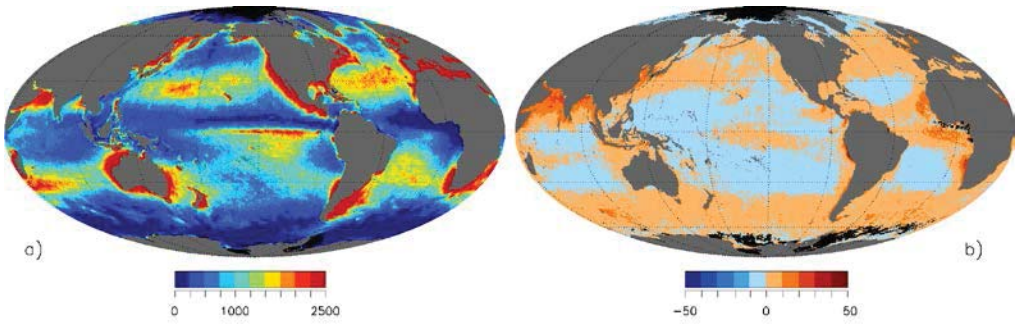


Figure 6.6: a) Total number of match-ups between MODIS and SeaWiFS (2003-2007) on a 3rd-degree grid and b) Mean relative difference ψ* between SeaWiFS and MODIS.

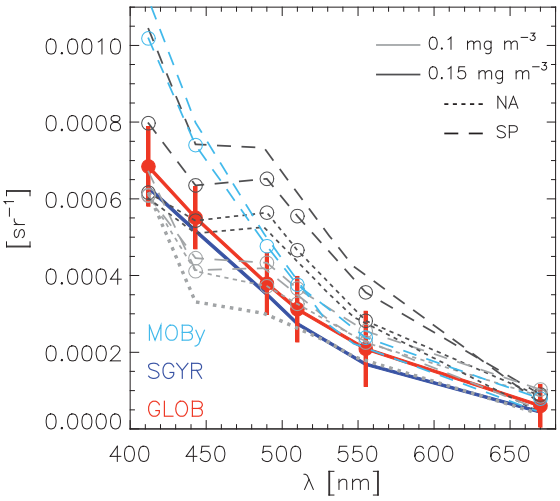


Figure 6.7: Uncertainty estimate σ obtained by co-location between SeaWiFS and MODIS, with global average (red) with standard deviation, and average over the subtropical gyres (blue). In grey are overplotted estimates of the uncertainty term proposed in [51] for SeaWiFS and MODIS as a function of Chla for the North Atlantic (NA) and South Pacific (SP) subtropical gyres. Validation results $\Delta_u$ obtained with MOBy data are shown in light blue. Curves with circles are for SeaWiFS.

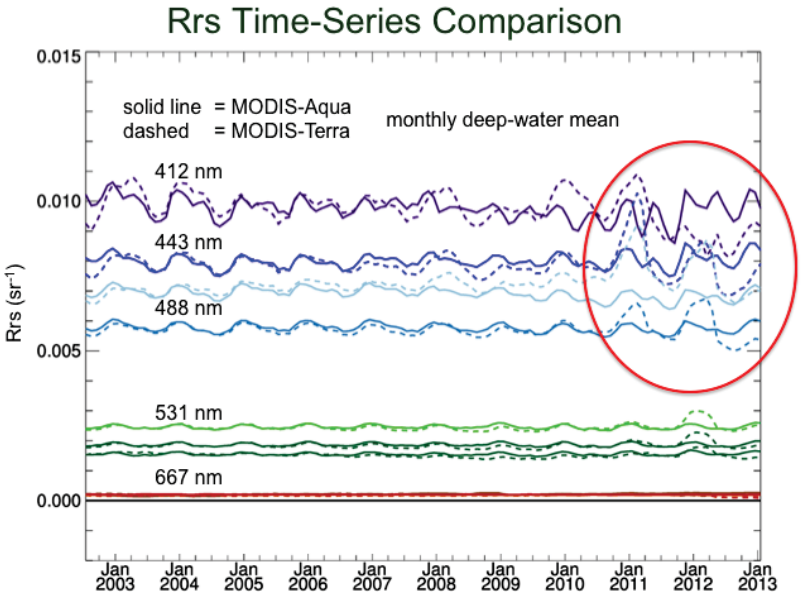


Figure 6.8: Comparative common-bin time-series of MODIS-Terra and MODIS-Aqua.

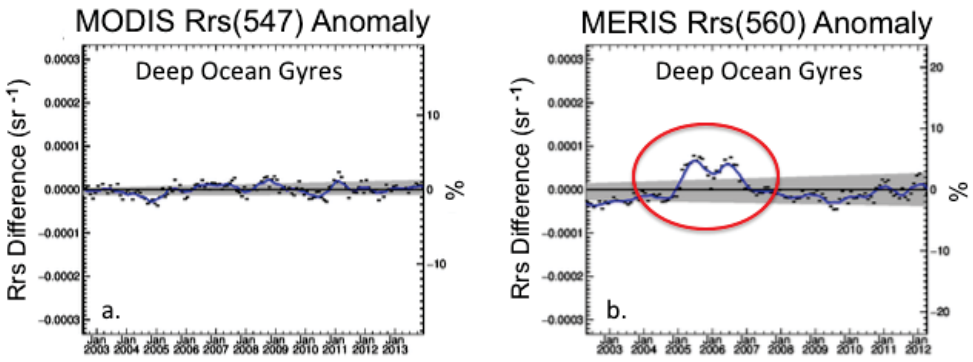


Figure 6.9: $R_{rs}$ anomaly analysis for a global clear water region.

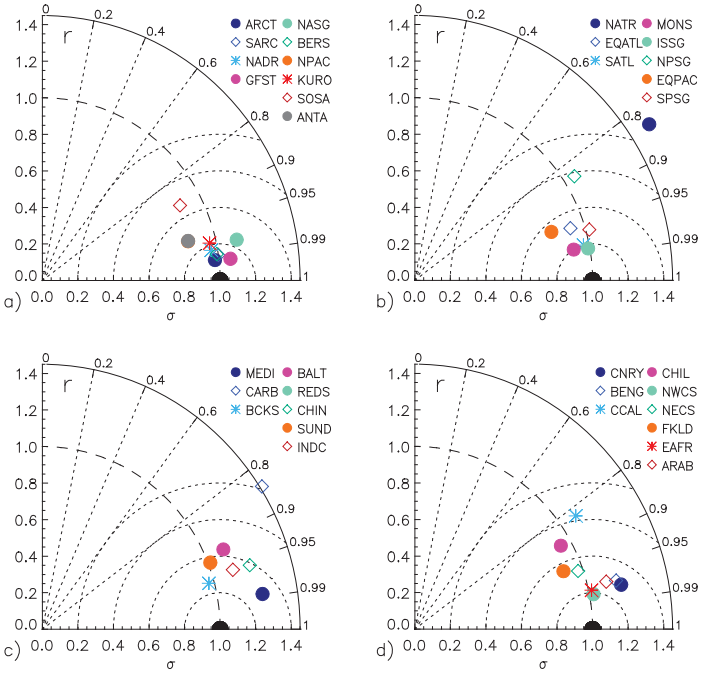


Figure 6.10: Taylor plots comparing SeaWiFS (taken as a reference) and MODIS-Aqua Chla time series averaged over biogeographic provinces associated with a) mid-latitudes, b) subtropical regions, c) marginal seas, and d) shelves and upwelling regions. See Longhurst [81] for province acronyms.

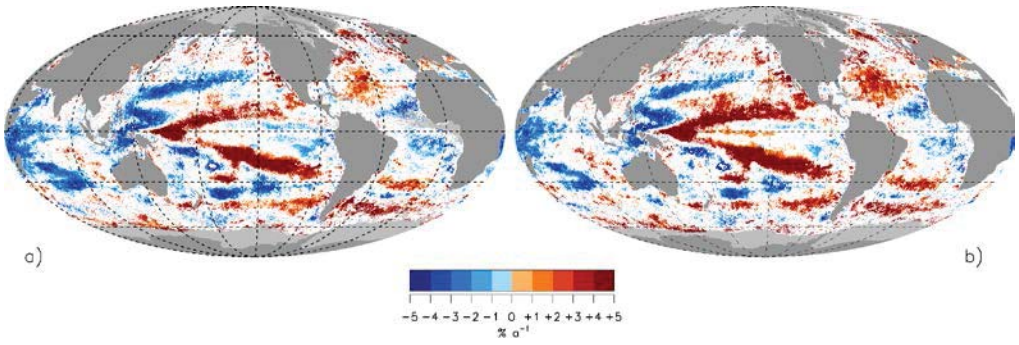


Figure 6.11: Linear trend for Chla (in % per annum) found over the period August 2002 to July 2011 for a) MODIS-Aqua and b) MERIS.